Journal of Economic Behavior & Organization xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

The under-appreciated drive for sense-making

Nick Chater[a,*], George Loewenstein[b]

[a] Behavioural Science Group, Warwick Business School, University of Warwick, Coventry CV4 7AL, UK
[b] Department of Social and Decision Sciences, Carnegie Mellon University, 5000 Forbes Avenue, BP 208, Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:
Received 14 October 2015
Accepted 28 October 2015
Available online xxx

Keywords:
Sense-making
Drive
Motivation
Simplicity
Information avoidance
Confirmation bias

ABSTRACT

This paper draws attention to a powerful human motive that has not yet been incorporated into economics: the desire to make sense of our immediate experience, our life, and our world. We propose that evolution has produced a 'drive for sense-making' which motivates people to gather, attend to, and process information in a fashion that augments, and complements, autonomous sense-making. A large fraction of autonomous cognitive processes are devoted to making sense of the information we acquire: and they do this by seeking simple descriptions of the world. In some situations, however, autonomous information processing alone is inadequate to transform disparate information into simple representations, in which case, we argue, the drive for sense-making directs our attention and can lead us to seek out additional information. We propose a theoretical model of sense-making and of how it is traded off against other goals. We show that the drive for sense-making can help to make sense of a wide range of disparate phenomena, including curiosity, boredom, 'flow', confirmation bias and information avoidance, esthetics (both in art and in science), why we care about others' beliefs, the importance of narrative and the role of 'the good life' in human decision making.

1. Introduction

In his 1864 masterpiece, *Utilitarianism*, John Stuart Mill embraced Bentham's Utility Principle, but proposed a more expansive conception of utility than the purely pleasure and pain-based concept proposed by Bentham. With his famous statement that "it is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied," (260) Mill drew attention to the importance of insight and wisdom, and of appreciating 'higher' pleasures such as art and music. Insight and knowledge, Mill felt, were goals, above and beyond the seeking of pleasure and avoidance of pain, that people should and do strive for. Yet subsequent implementations of different conceptions of utility in economic theory have rarely dealt with the types of 'higher' pleasures discussed by Mill in a substantive fashion. Even the broadest notions of utility that have been proposed, for example 'ego utility' or belief-based utility, fail to account for the enormous time, money and attentional resources that people devote to sense-making.

In this paper we posit the existence of a 'drive for sense-making' which, we argue, is analogous to better known drives such as hunger, thirst and sex. We review diverse research on sense-making from psychology, then lay out the basic elements of a theoretical model of utility maximization that incorporates sense-making as an ingredient of utility. In our model, individuals have two goals that drive their behavior: (1) to construe our lives in a positive fashion (valence); and (2) to

* Corresponding author.
E-mail addresses: nick.chater@wbs.ac.uk (N. Chater), gl20@andrew.cmu.edu (G. Loewenstein).

http://dx.doi.org/10.1016/j.jebo.2015.10.016

construe our lives in a way that makes sense (sense-making). Given that both involve events occurring purely in the mind, they might seem to favor a life consisting of purely mental activity—e.g., fantasizing that one's life is desirable and makes sense. However, both of these tendencies are severely constrained by the brain's autonomous sense-making capabilities, which are involuntary and have the sole objective of maximally simplifying information (i.e., are not influenced by valence). We can make deliberate (non-autonomous) decisions about *whether* to collect information, but we have no ability to dictate how we will interpret the information we collect, nor, of course, to choose what we will discover when we choose to collect information.

The model has novel implications both for when people choose to obtain or avoid information, but also for more ordinary economic decisions. Given the constraints on self-deception inherent in the autonomous processes' sole concern with sense-making, in most situations the safest and most direct route toward ensuring that our image of our lives is favorable and makes sense, is to live a life that has positive features and that can be made sense of. We discuss a wide range of applications of the model, including curiosity (the desire for information for its own sake), boredom, flow (the pleasure of sense-making), confirmation bias, information avoidance, consumer choice, esthetic preferences (including both art and science), concern about others' beliefs, conspiracy theories and religion, the importance of narrative, and the role of the 'good life' in decision making and life satisfaction.

2. Sense-making and simplification

The Oxford English dictionary defines a drive as "an innate, biologically determined urge to attain a goal or satisfy a need" (Stevenson, 2010, p. 535), and this is the sense in which we use the term here. Due to profound limitations on how much information the brain can process and store, as well as the desire for efficiency in communication, evolution has produced elaborate neural mechanisms for the simplification and distillation of information. Such processes guide perception, language, memory, and a wide range of other cognitive processes. Knowing that the object in front of one is a table, for example, we can safely assume that it is solid, flat, elevated from the ground, can hold a laptop and drinks, as well as myriad other properties that are difficult to enumerate because they are so fully assimilated in our mental representations of tables that we are unaware of their existence.

The view that perception and cognition seeks to make sense of the world has a long and varied history. For example, Gestalt psychology, a school of psychology that thrived in the early 20th century, was concerned with the acquisition of meaningful perceptions in a chaotic world. Gestalt psychologists enumerated a series of 'laws' or 'principles' dictating how the mind makes sense of the environment by constructing global wholes—'Gestalts'—from otherwise chaotic stimuli (Koffka, 2013/1935; Rock and Palmer, 1990).

Fig. 1 shows some classic stimuli that demonstrate Gestalt principles. In the left hand stimulus (due to the celebrated Italian psychologist Kanizsa, 1979), postulating an invisible square that covers some of the black blobs 'makes sense' of the missing elements. Similarly, the central stimulus can best be made sense of by postulating a 3D white 'wire frame' cube which partially occludes the black circles seen as behind it. The integrated nature of this interpretation is made particularly evident in virtue of the ambiguity of the wire-frame cube—it is a so-called Necker cube, much discussed psychology and neuroscience). When the cube 'flips' from appearing to be viewed from above and tilted to the viewer's left, to appearing to be viewed from below and tilted to the viewer's right, the black circles at the vertices of the cube correspondingly appear to change 'depth.' The third figure, Idesawa's (1991) sphere, is perhaps even more remarkable. The brain creates a smooth white sphere (which appears, to many observers, to be a brighter white than the surround) radiating conical black spines in three dimensions, from a collection of flat black geometric shapes. According to Gestalt theory (e.g., Chater, 1996; Pomerantz

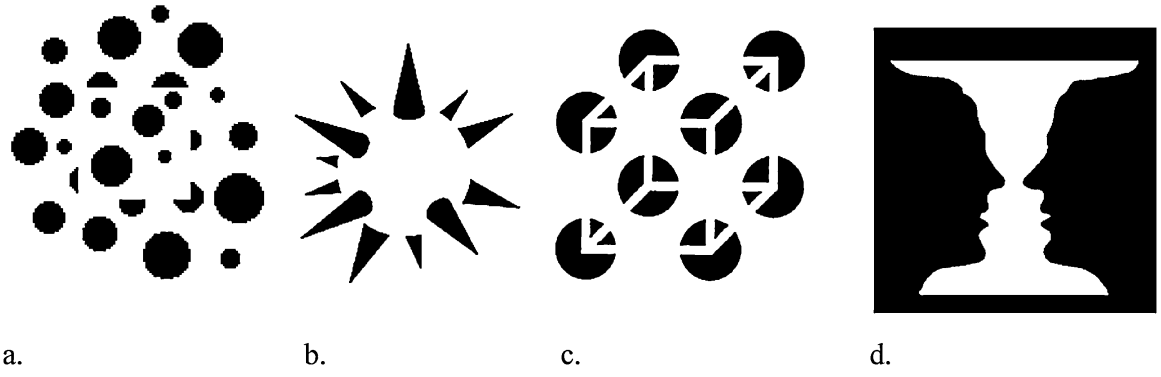

Fig. 1. Sense-making in perception. The brain prefers organizations which provide a simple encoding of the sensory input. (a) A white square that the brain creates in order efficiently to encode the various 'lost' chunks of some of the filled black circles (figure devised by Italian psychologist Gaetano Kanizsa); (b) a white sphere, with projecting black 'spikes' encodes this otherwise haphazard set of roughly triangular 2D shapes (figure devised by Japanese vision scientist Masonori Idesawa); (c) a virtual white 'wire-frame' cube encodes the 2D pattern of missing lines on the filled black circles; (d) the face-vase illusion, devised a century ago by Danish psychologist, Edgar Rubin.

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

and Kubovy, 1986), the sphere-with-spines is constructed because it provides a simple explanation of the precise outlines and locations of the black shapes. Finally, the well-known Rubin vase, or face-vase stimulus (which can be perceived either as a vase or as two faces in profile), illustrates how the elements of a stimulus can be interpreted very differently, depending on the 'whole' into which they are assimilated. Flipping between the face and vase interpretations involves changing which part of the image is 'figure' and which is 'ground.' Given that the brain encodes figures in terms of convex 'bumps' rather than concave 'hollows' (Hoffman and Richards, 1984), the features encoded in the face interpretation (nose, lips) are, on the vase interpretation, not even encoded as parts of the stimulus at all, but as lying between these features. This stimulus is a classic illustration of different concepts from Gestalt psychology, including the idea (embodied in our model) that the brain can only make sense of the world in one way at any point in time—that is, we see the face or the vase, but not both.

If the brain has a drive for sense-making, then it presumably requires some measure of *how much* sense different interpretations make, whether these are interpretations of perceptual stimuli (see Fig. 1), or more abstract aspects of the environment. Parallel traditions in philosophy (Kemeny, 1953), psychology (Hochberg and McAlister, 1953), neuroscience (Blakemore, 1993) and statistics (Rissanen, 1987 and Wallace and Freeman, 1987) have suggested that explanations can be viewed as *encoding* the data they seek to explain; that the complexity of an explanation, for a particular set of data, can be measured by the length of the resulting code.

To get an intuitive sense of how the approach works, imagine our encoding language is English, and suppose that one person is attempting to describe the images in Fig. 1 with sufficient precision that another person is able to recreate them. So, in Fig. 1, high level descriptions, such as 'there is a field of different sized black blobs, with a pure white square in front,' 'the eight vertices of a white wire frame cube are at the center of equal-sized black circles,' or 'a perfectly white sphere is radiating black thorn-like cones in all directions,' are likely to be much more useful than an attempt to describe the precise outlines and locations of 2D black shapes. Of course, for the receiver of the message, the detail needs to be specified too; but it will be much more efficient to ask about the size and location of the white square, the wire frame or the spines, rather than focusing at the level of 2D outlines. The essence of the idea is that interpreting data by imposing some sense or structure helps the brain encode data efficiently; so we can invert this logic, and use the brevity of the encoding using that interpretation as a measure of the *amount of sense* that it makes of the data. It is then natural to assume that, other things being equal, the brain will prefer the briefest explanation that it can find.

Though useful in driving intuitions, natural language is a poor choice for coding for a variety of reasons, most obviously because of its imprecision and ambiguity, and less obviously because the very idea of code lengths in natural language leads to paradoxes (e.g., the Berry paradox, Chaitin, 1995). To avoid these problems, it turns out that we can instead use a *programming language*, in the sense used in computer science: such languages are precise, unambiguous and paradox-free; and appear, in principle, rich enough to express any well-defined interpretation or structure. Indeed, the idea of measuring the complexity of explanations by codes in a programming language is the core of a rich mathematical theory, Kolmogorov complexity (Li and Vitányi, 2009), which has a variety of psychological applications (e.g., Chater, 1999 and Chater and Vitányi, 2003, 2007). In this paper, though, we leave these technical issues concerning measures of code-lengths aside. We focus instead on modeling how sense-making can function as a drive. Here, we merely note that such quantification of sense-making is both feasible and has been widely explored in prior work.

We have noted that the brain attempts to make sense of the world; that sense-making may be measured by the simplicity of the explanation of the data, and that simplicity may, in turn, be measured by code-length. Yet why should there be a drive for sense-making, as we argue in this paper? Indeed, why should sense-making have any affective 'valence' or motivational force at all? Affective states are typically viewed as involved with motivation: we seek pleasant, and avoid unpleasant, sensations. But sense-making appears to operate autonomously, irrespective of conscious control or motivational state. Indeed, the process of understanding the physical and social world is, to a large degree at least, a reflex (Fodor, 1983). Our visual and auditory systems automatically make sense of sensory input; we cannot decide whether or not to understand what people are saying, or whether or not to go through the complex inferential processes required to recover their intentions and motives (e.g., Clark, 1996; Levinson, 2000), whether or not to appreciate the significance of what we are told, and so on. Indeed, any lack of autonomy for such processes could be highly dangerous: if we had the ability to "see what we want to see," in a literal, rather than a metaphorical sense, this would appear entirely to undermine the informational gathering role of the senses. But to the extent that sense-making is autonomous, we might expect it to be affectively neutral—we do not need seem to require motivation to make sense of the world any more than we need motivation to maintain our balance or to blink when an object rapidly approaches our eyes.

We suggest that a drive for sense-making has an analogous function to that of other drives. Note that many internal, and automatic, biological processes have behavioral extensions. For example, temperature regulation in the human body relies on a range of internal mechanisms such as blood flow, piloerection, sweating, and so forth. There is no need to 'feel hot' or 'feel cold' in order to trigger changes in blood flow or sweat production—indeed such changes are not under conscious control. However, people sometimes find (or propel) themselves into environments in which internal processes are inadequate for thermoregulation. In such situations, evolution has produced motivational mechanisms to induce us to take actions, either to reduce an aversive affective state or to enhance a positive one. Far from being automatic, such behaviors require deliberation: we decide to turn on an air conditioner or put on a sweater or cover ourselves with a blanket. And our motivations concerning temperature can compete with other motivations, such as saving money on fuel bills or wanting to look stylish. Similarly, the drives associated with hunger and thirst motivate us to acquire and consume the "raw materials" upon which automatic processes of digestion and absorption operate; the occasional, but sometimes overwhelmingly powerful, drive to be able

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

to breathe will provide powerful motivation to come up for air when we are diving or use an inhaler if we are suffering an asthmatic attack, complementing autonomous respiratory processes.

We suggest, then, that sense-making processes may be as automatic as digestion and respiration, but that, analogous to these other drives, in some situations the motivational system is required in order to provide the appropriate materials upon which these processes can operate. Although most sense-making mechanisms operate autonomously, without conscious awareness or deliberate direction, in some situations, information processing alone is inadequate to transform disparate information into simpler representations. In such situations, we argue, our brains have evolved mechanisms to motivate us to gather, attend to, and process information in a fashion that augments, and complements, autonomous sense-making. Such actions can take the form of deliberate internal processes, such as searching memory or directing attention, or making minimal actions, such as shifting one's gaze, or actions varying in complexity from web searches or to taking a course, hiring a private detective or funding a research program. In short, we have a drive for sense-making because we frequently need to direct and "feed" our automatic sense-making machinery appropriately.

Most drives, furthermore, operate via a carrot and stick. The stick is the aversive feeling which arises when the drive is not met. One feels uncomfortable, for example, when one's body temperature rises above, or falls below, the 98.6° F set-point. The carrot is the pleasure of satisfying the drive, which intensifies with the strength of the drive itself, a pattern that psychologists call 'aliesthesia'. Continuing with the temperature example, anything that helps to restore our body's temperature set-point, like putting one's hand in cold water on a hot day, or in hot water on a cold day, feels good. The drive for sense-making fits such a pattern: The inability to make sense of stimuli, or for that matter of one's life, is generally aversive, and sense-making is especially pleasurable when the drive for sense-making has been activated. Thus, for example, it is much more pleasurable to get the answer to a question if the question is asked first, with a pause for one to come up with the answer, than if the question and answer are delivered at the same moment in time (Hsee and Ruan, 2014).

Different drives have different triggers, some involving internal bodily states and some involving external stimuli. For example, the strength of the sex drive at any point in time depends both on the body's state (e.g., hormone levels, which, in turn can depend on time of day, time of the month, and time since last sex), and external—e.g., visual—stimuli. The drive for sense-making is, again, not an exception; its activation depends both on internal (cognitive) states and external, informational, stimuli.

2.1. Four key features of sense-making

Several observations about sense-making and its connection to simplification provide the foundation for our basic theoretical framework. The first is that both sense and sense-making are pleasurable (and their opposites are aversive). So, for example, the stimuli in Fig. 1 may give us (an admittedly modest amount of) pleasure because we have can make sense of them successfully; and, in addition, the experience of sense-making is itself pleasurable: We may experience a momentary positive 'frisson' when we suddenly 'find' an elegant interpretation (the white square, the wire-frame cube, the white, spiny, sphere, or the 'faces in the vase'). Thus, as with contemporary theories of decision making which posit that people derive utility both from outcome states and from changes (or deviations from reference points), sense-making is a function of both levels (sense) and changes in the sense one has made (sense-making). Pleasure from sense captures, for example, both the satisfaction of being able to tell a coherent story about our life, but also the satisfaction of gaining new information that leads to a refinement of that story. Ex ante, at least on average, we should always expect new information to enhance sense-making. However, this is not always the case. For example, a neat story one may have told about one's life could be disrupted by the arrival of new information, leading to a negative change in sense and the opposite of sense-making. Indeed, as we discuss below, fear of such order-disrupting information can lead to information avoidance and confirmation bias.

The second principle is that not all sense-making is equally pleasurable. Some types of sense-making are more *important* to us than others, and pleasure from sense-making (as well as pain from failures to make sense) depends critically on importance. For example, it is far more pleasurable to make sense of one's life than to solve a Sudoku puzzle or a crossword, and it is more pleasurable to make sense of one's own life than to make sense of someone else's. We gain still less pleasure from aimlessly glancing around the room, or staring vaguely out of the window.

Some pieces of information are more important to us than others not merely because of the 'interest' of the topic, but because they are a 'key' which allows us to provide a better explanation of many other things we may be interested in. So acquiring a small amount of new information might, in principle, substantially reduce the complexity of a wider body of information. For example, finally discovering the perpetrator of a whodunit one has been immersed in for weeks is typically more pleasurable than discovering the perpetrator at the end of a short-story, in part because the information brings coherence to a lot more information in the former case than in the latter. And, of course, both of these are more pleasurable than being told the identity of the criminal in a whodunit one has not read.

To account for such differences in our model, we assume that sense and sense-making is, in fact, pleasurable or painful to the extent that it does help us to make sense of aspects of our lives. Solving a Sudoku puzzle, by this assumption, will be less satisfying than finding an explanation for a friend's sudden coldness. Solving the Sudoku puzzle makes sense only the current stimulus, with few or no wider implications beyond the momentary situation. But a better understanding a friend's unexpected behavior may have implications for much broader aspects of one's life, such as the state of an important on-going relationship. When two forms of potential sense-making are equally important to us, however, in the sense of shedding similar light on one's life, then purely informational considerations will dictate their hedonic impact. Thus, for

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

example, purely informational considerations should be able to explain why we get more pleasure from solving one Sudoku puzzle than another, or from solving a Sudoku puzzle over playing a game of Tetris.

A third principle is that the pleasures and pains of sense and sense-making depend on expectations. Perhaps one distinctive feature of the stimuli in Fig. 1 is that the degree of sense-making possible is unexpected: e.g., normally 2D shapes and patches do not support such complex interpretations. In other situations, however, we might go into a situation with high expectations of sense-making, but find ourselves disappointed or frustrated. Prior research in diverse domains has found that pleasure and pain tends to result from comparison of attainments to reference points (Kahneman and Tversky, 1979), which are very often determined by expectations (Kőszegi and Rabin, 2006). We propose that sense-making is no exception. Such reference-dependence can explain why we find stimuli aversive when they are surprisingly difficult to make sense of (e.g., blurry images or text, mis-spellings, or even grammatical errors). Dependence on expectations appears to be required to capture the fact that, in a museum, we derive no pleasure or pain from the blank walls between paintings because we don't have any expectations that sense is to be found in these expanses. Likewise, the feeling that there is sense to be made, but we are unable to make it, can be agonizing, as exemplified by people who fruitlessly ruminate about key events in their lives, in a futile effort to make sense of them.

A fourth principle is that the pleasures and pains of sense-making are not symmetric. Consistent with the notion of loss aversion (Kahneman and Tversky, 1979) and the more general phenomenon that the bad tends to be stronger than the good (Baumeister et al., 2001), negative deviations from expectations tend to be more aversive than positive deviations are pleasant. Thus, believing one has made sense of a situation, but then discovering one has not, is an overall aversive experience, despite the fact that one is left in the same situation in which one began. Likewise, receiving information that challenges the sense one has made of the world, but then having that information discredited, leaves one worse off (albeit partly due to new appreciation of the fragility of the sense one has made) than one was at the outset.

3. A simple model of sense-making

We suppose that there is a psychological distinction between knowledge about aspects of our lives, K, that we are at any given moment attempting to explain; and general background knowledge, B, e.g., about the physical and social world. That is, K, represents the information that the brain is currently trying to make sense of. This boundary is, as we shall see, flexible. At one moment, we may be focusing purely on the interpretation of a visual stimulus, such as those in Fig. 1, so that K may be limited to current perceptual input; at another moment, we may reflect on how our life, or our relationships, are progressing, so that the knowledge that we are attempting to explain is far broader. We suggest, though, that we are only able explain some subset of our knowledge of our lives, K, at any moment; we treat anything else that we know as part of our background knowledge, B.[1]

We propose that people form the explanation, $e(K|B)$, of the to-be-explained knowledge, K, which best fits that knowledge, given the background knowledge B. We assume that the process of inferring this explanation is independent of their *evaluations* of their lives, just as it is typically assumed in perception that the favored perceptual interpretation is autonomous and hence independent of the goals of an agent (e.g., Fodor, 1983; Pylyshyn, 1999).[2] For clarity, and in line with the observation concerning perceptual organization that people can only entertain a single explanation of a sensory input (e.g., the face/vase of Fig. 1) or of their life at any moment, we assume that $e(K|B)$ corresponds to the single simplest explanation or construal of K, rather than a probability distribution over construals.

Thus,

$$e(K|B) = \underset{e'}{arg\,\min}\, C\left(e', K|B\right), \tag{1}$$

where $C(e', K|B)$ is the code-length for encoding knowledge, K, using explanation e' based on background knowledge B. This term itself can be expressed as the sum of the code-length of the explanation e' itself, $C(e'|B)$, and the code-length required to specify knowledge K given the explanation, $C(K|e', B)$. That is, $C(e', K|B) = C(e'|B) + C(K|e', B)$. A preferred explanation will be itself simple (as measured by the code-length $C(e'|B)$) but also accurate, hence providing a simple code for the knowledge the individual seeks to encode, as measured by $C(K|e', B)$. This is a foundational assumption of our model of sense-making: Given background knowledge, the autonomous process prefers the simplest explanation of the to-be-explained information, K, that it can find. This is a difficult optimization problem—indeed, in general it is uncomputable; that is, no computable

[1] As our attentional focus shifts, the boundary between the information that we are attempt to explain, K, and the background knowledge that we presuppose, B, may change in important ways. For example, while we may momentarily be attempting to make sense of the immediate actions of the characters in our work of literature, film or play (so that the relevant K, in the moment, is quite narrow), it may be that when we brought them our attention to think about our own life, the people around us, or our society (i.e., we now shift to a much broader K), the analysis of the work of art has changed our interpretation of these much broader aspects of our lives. As we note below, the degree to which an experience or a work of art can be made sense of in and of itself, or, rather, changes how we make sense of other aspects of our lives, may be important in distinguishing experiences which we take to be "deep" from those we consider to be "shallow" (e.g., solving a Sudoku problem presumably has no wider implications beyond the stimulus itself).

[2] Although a useful simplification, this is may be unrealistic. There is a wealth of research in psychology that appears to show that even our most basic perceptions are affected by motivational factors such as drives, goals and desires although the interpretation of such research is controversial.

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

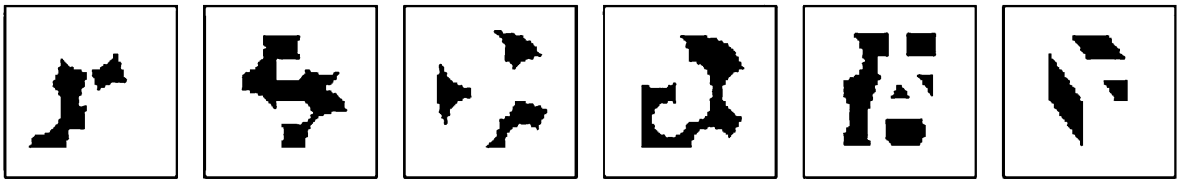

Fig. 2. Stimuli which are initially puzzling, but which are actually surprisingly familiar.

Source: Fukushima (2001).

process can map sets of data into their shortest descriptions. But this is not unexpected—the problem of making best sense of our environment is inevitably an open-ended problem, closely related to creating the best possible science.

Now, suppose that we learn a new piece of information that might potentially modify our construal of the knowledge, K, that we are trying to make sense of. We model this by adding a new piece of information, a, to our background knowledge, which then becomes $B \cup \{a\}$. Now, our best construal of our life, in the light of a, written $e(K|B \cup \{a\})$, is:

$$e(K|B \cup \{a\}) = \arg\min_{e'} C(e', K|B \cup \{a\}) \tag{2}$$

It could, of course, be that a is irrelevant to making sense of what we know of our life (the knowledge, K), and so the best choice of construal remains unchanged (i.e., $e(K|B \cup \{a\}) = e(K|B)$). But it is also possible that a leads to some updating of the best construal. For example, getting some positive (or negative) teaching feedback may make us re-construe past, perhaps somewhat ambiguous, remarks about our teaching performance; learning a medical fact might require us to reinterpret some symptoms we have been experiencing; and so on.

Just as in science, a single piece of information (a crucial observation or experiment) can change our current choice of best 'theory.' Indeed, even without any new information, we may find ourselves 'flipping' between two roughly equally good interpretations. We might, for example, oscillate from feeling genuine confidence in our abilities and suspicion that other people are giving us apparently positive feedback only out of kindness; between believing that our company strategy will win new customers and fearing that it will fail hopelessly; and so on. In such cases, there may be little evidence available to decide between interpretations.

This type of 'Gestalt switch' is illustrated in the Necker Cube (Fig. 1, where we alternately interpret the cube as tilted downwards to the left or upwards to the right) and face-vast illusion (also in Fig. 1); and it often discussed in the philosophy of science (e.g., Kuhn, 1962, Scheffler, 1972; Wright, 1992). For this reason, the best explanation is not, strictly, a *function* of the relevant knowledge to be explained, K, and background knowledge, B, because the same knowledge can lead to different interpretations (whereas a function, of course, can only map one input onto a single output). Nonetheless, purely to keep our notation simple below, we will ignore the possibility of such spontaneous 'flips' and suppose that the brain chooses a single best construal or explanation e of knowledge, K, given background B, so that we will write the best construal e as a function $e(K|B)$ of the relevant knowledge.

Now, on learning a, the complexity of our new best explanation ($C(e', K|B \cup \{a\})$) may be more or less than the complexity of our previous explanation, $C(e', K|B)$. For example, suppose that K concerns relevant aspect of our visual input, when looking at the stimuli shown in Fig. 2.

You may have spotted the hidden pattern (take a moment to see if you can find it). But if not, considering the additional piece of information a = "these are the first six letters of the alphabet, in block capitals." With this information in mind, these patterns suddenly become easier to make sense of. And, of course, according to this interpretation, the code for the stimuli, K, will be far shorter. For example, imagine trying to convey these shapes by a written description: once we have the 'key' that these are block capital letters, with various occlusions, it becomes possible to encode them reasonably briefly, by describing the color, thickness, and font of the letters and then describing the sizes of centers of the 'white' circles, lines, and so on, that overlay them (notice to express this code, we need to draw on background knowledge, B, about the letters of the alphabet, fonts, colors, shapes, and so on). By contrast, describing them purely as highly irregular patterns will require an extremely long and convoluted description. So, with the 'hint' a, that the stimulus is a sequence of letters, the code for K becomes far simpler, and hence makes far more sense: that is, our estimate of $C(e', K|B \cup \{a\})$ is much less than our estimate of $C(e', K|B)$)—only with the hint can we make sense of the stimulus.[3]

While adding new information, a, to our background knowledge, B, typically makes our encoding of the data more efficient, the opposite is also possible. We might, for example, be told that the patterns in Fig. 2 are markings on a tablet from Ancient Egypt, or images picked up from a scan of the sea-bed. If these assertions were sufficiently credible, then we would need to

[3] Indeed, in this case, a stronger condition may be true: that the best explanation e', for K, may actually *include* the 'additional' information a. If so, by choosing the best explanation, we automatically postulate that a is true—in principle at least, it does not need to be given as a hint: it is 'implicit' within the (best explanation of the) data itself. Of course, even in this situation, the hint may be required for us to *find* the explanation. So, when looking at Fig. 2, the hint may not be required for us to realize that panels represent the first six letters of the alphabet. This stronger condition is of particular interest in perception (e.g., Chater, 2005), but is not our focus below.

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

jettison our apparently efficient encoding (though it might still help us remember the patterns), because it is inconsistent with our background knowledge (there was no Latin alphabet in Ancient Egypt, and so on).

How, does simplicity relate to sense-making? We propose that the amount of sense that we make of knowledge about some aspect our lives, K, or any other body of information, is inversely related to the complexity of the explanation we are able to create. The brain is continually seeking the simplest explanation, and hence the explanation corresponding to the shortest code.

$$S(e(K|B) = f(-e, K|B)), \tag{3}$$

where f is a strictly positive monotonic function.[4] It follows immediately that the explanation that minimizes code length, is the same as the code which maximizes the degree of sense-making:

$$e(K|B) = \underset{e'}{arg\,\max}\ S(e', K|B) = \underset{e'}{arg\,\min}\ C(e', K|B) \tag{4}$$

Here, we have developed a model of how the brain attempts to find the explanation to make the most sense of given information, whether a momentary sensory input, or our knowledge of our entire life, in light of background information. Sense-making is related to the simplicity by which the given information can be encoded—the brain seeks the simplest explanation that it can. If we have a drive for sense-making, then it is natural to conjecture, as we have noted above, that sense-making, i.e., finding simpler explanations, should be pleasurable; and the opposite should be aversive. But, of course, when we explain our lives, or encounter new information that might modify such explanations, we are not concerned only with the degree of sense-making; we are also concerned with what the current sense we have made, $e(K|B)$, says about how our life is going. Sense-making might, in principle, lead us to a wonderfully compact explanation of our lives as consisting of relentless calamity. In short, sense-making is not our *only* drive or motivation. In the next section we consider how sense-making can be combined with other factors—lumping together all such other factors into a single quantity of utility, for analytical tractability.

3.1. Sense-making and utility

An explanation, $e(K|B)$, of our lives may paint our lives in a good or a poor light and may suggest a rosy or gloomy future in the light of the aspects of our life, K, that we are thinking about. Of course, if K refers to something extremely specific, such as the immediate sensory input, then the explanation $e(K|B)$ may have no significant implications for our lives at all. But often the information, K, that we are trying to explain will concern our own behavior or performance, our relationships with significant people in our lives, and so on. In these cases, which explanation we settle on may have substantial implications for our evaluation of our lives. Let us call a person's evaluation of how positively (or negatively) some aspect of life, K, is going, $V(e(K|B))$, paralleling the term introduced above for much sense they can make of their life, $S(e(K|B))$.

We assume that individuals pursue two corresponding goals: (1) to view their world and their life as positively as possible, and (2) to make maximal sense of their lives. We assume that a person's "absolute" utility, $U_A(e|K, B)$, depends on the absolute valuation of their life and the absolute amount of sense they can make of it:

$$U_A(e(K|B)) = +U_A(V(e(K|B)), S(e(K|B))) \tag{5}$$

We make the further natural assumption that $U_A(e(K|B))$ is increasing in both arguments.

But absolute utility, U_A, is only part of the story. Following Kőszegi and Rabin (2006), we allow that people's overall utility depends both on absolute utility and on how the components of this utility depart from expectations. That is, we assume that people care separately about whether their construal, $e(K|B)$ of their life is becoming more or less positive; and whether it implies that their life is making more or less sense.

If changes relative to expectations are key, as in the Kőszegi-Rabin framework, then a person will care about the degree to which, after learning new information a (so that her background knowledge is now $B \cup \{a\}$), her life is better or worse than its prior value, given her previous background knowledge state B. We call this reference-dependent value term, V_R, and write:

$$V_R(e(K|B \cup \{a\})) = \mu_V(V(e(K|B \cup \{a\})) - V(e(K|B))), \tag{6}$$

where μ_V is a loss-averse value function of the type used in prospect theory (i.e., concave in gains, convex in losses, and where losses loom larger than the corresponding gains, Kahneman and Tversky, 1979). Similarly, a reference-dependent sense-making term, S_R, i.e., the change in the degree to which life makes sense in the light of a, can be written:

$$S_R(e(K|B \cup \{a\})) = \mu_S(S(e(K|B \cup \{a\})) - S(e(K|B))), \tag{7}$$

where μ_S is similar a loss-averse value function of the same general form. Loss aversion has, of course, important implications for the impact of information. For example, if a brings bad news concerning one's evaluation of how one's life is going (e.g.,

[4] Below we do not place any explanatory burden on the functional form of f, so we could take it to be the identity function. It is possible that a more complex form for f may be theoretically useful in some contexts.

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

poor teaching ratings, a failed grant application, or a share price collapse) and/or makes one's life make less sense (e.g., one learns that a cherished story about one's childhood is, in fact, probably a false memory), then one's absolute utility will be reduced. That is, our absolutely utility will fall from $U_A(e(K|B))$ to a lower value $U_A(e(K|B \cup a))$. But if one subsequently discovers that a is entirely false—e.g., the person from whom we overheard this information was actually talking about someone else entirely, other things being equal, the absolute utility term will now return to its previous value, $U_A(e(K|B))$.[5] However, due to loss aversion, the changes in the reference-dependent value and sense-making terms, V_R and S_R will not 'cancel out' when a is first added and then rescinded. Whether a increases or decreases value or sense, its being discredited will always leave one worse off than prior to receipt of the information, because, due to loss aversion, the negative change will always outweigh the positive one. This fits, for example, our intuition that experiencing a fear which then turns out to be unfounded is, overall, a negative experience, or that hearing good news which turns out to be illusory, is also unpleasant overall.

Valuation and sense-making will often move in opposite directions upon the receipt of new information. For example, news of one's child's arrest for narcotics possession might help to make sense of a wide range of otherwise incomprehensible information (e.g., why his moods were so erratic), thus raising S, but would certainly have a negative impact on V. In contrast, learning that someone you suspected might dislike you had dedicated a book to you might fill you with pleasure, yet disrupt the neat sense you had made of your relationship with the individual. The fact that information can often have opposite effects on sense-making and value might also help to explain the common observation that, after receiving long-anticipated news that turns out to be bad, we can sometimes have a strong feeling of relief: Perhaps this feeling arises because, though the outcome is not what we had hoped for, the certainty of our knowledge puts us back in a position to make sense of our lives. More broadly, such 'bitter-sweet' reactions, which seem a ubiquitous part of human experience, seem to require explanation in terms of potentially antagonistic forces, such as optimizing V and S, and are difficult to understand in any unitary account (e.g., where we are interested only in optimizing V).

Overall or total utility, U_T, in the light of new information a, is an increasing function of three quantities: the absolute utility of the state when a is known (itself depending on valuation and sense-making), and the two reference-dependent terms:

$$U_T(K, a|B) = +U_T\{U_A(e(K|B \cup \{a\})), V_R(e(K|B \cup \{a\})), S_R(e(K|B \cup \{a\}))\}, \tag{8}$$

or, more compactly,

$$U_T(a) = +U_T\{U_A(a), V_R(a), S_R(a)\}, \tag{9}$$

where $U_T(a)$, the total utility after learning new information a, is assumed to be increasing in each of the three arguments. For simplicity, we can, though we need not, assume these utility terms are additive, so that our total utility term U_T, in the light of data a, can be written:

$$U_T(a) = \alpha U_A(a) + \beta V_R(a) + \gamma S_R(a), \tag{10}$$

where α, β, and γ are constants. Here, the "reference" level for valence and sense-making, $V(e(K))$ and $S(e(K))$ is determined before the new information a is known.

In the model just described, as a result of reference-dependence, expectations matter critically. For example, we should find it aversive, if we expected to finally learn the identity of the criminal in the last episode of an extended TV series, to discover that the anticipated information is not forthcoming (perhaps because the producers decided to add a second season). If the final episode did promise to reveal the perpetrator, moreover, but our cable connection failed at the crucial moment of revelation, our frustration would likely be even more profound.

4. Applications

In prior sections, we outlined an account of a drive for sense-making, explained how an account of sense-making can be integrated into existing accounts of utility and decision making, and finally explored how sense-making may be important in understanding how we assess the value of new information. In this section, we show how the drive for sense-making, and the specific theoretical perspective just outlined, can help to make sense of a wide range of disparate phenomena.

4.1. Curiosity

Curiosity is, perhaps, the most obvious application of our theoretical framework. Curiosity, by definition, refers to intrinsically motivated seeking after information—i.e., in the absence of anticipated material payoffs, or more broadly, where the information we are 'curious' about has no implications for our evaluation V of our lives. In our framework, curiosity can be easily understood as a manifestation of the drive for sense-making.

Curiosity in our framework could arise in two situations. In the first, an individual might get new information but be unable to make sense of it. In that case, curiosity would be focused on the desire for information that could make sense

[5] Of course there might be secondary effects, such as admonishing oneself for jumping to conclusions.

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

of the new information. Thus, for example, the website 'Upworthy' presents visitors with a series of headlines such as "If You Have To Tell Your Kids This Stuff, Then You Probably Aren't A White Person," and "Some Think It's Just Plain Kinky, But The Amount Of Trust This Lifestyle Takes Is Staggering" which evoke curiosity by presenting intriguing but confusing information that, the visitor is implicitly promised, can be made sense of by clicking on the link.

In the second case, curiousity is triggered by the perception that new information could help to make sense of existing, stored, information. A scientist, for example, would be tremendously curious to learn about a new theory that could make sense of otherwise disparate, unorganized data in his or her field, and would also be curious to obtain new information which, integrated with the existing data, might result in such a comprehensive understanding. Similarly, a person may research their family history in the hope of 'making sense' of their lives; read about 19th century Russian history to try to 'make better sense' of much loved Russian novels; or pore over a map of a location where they have been on a walking holiday, to understand the terrain. Note, too, that curiosity about fictional stories, which may be powerful enough to keep us reading or watching for many hours, can be understood in terms of the sense-making term, S, but fictional events typically do not, of course, have direct implication for our valuation, V, of our own lives.

A literature review by one of the authors (Loewenstein, 1994) proposed that curiosity derives for "an information gap" that becomes salient to the curious individual, and identified four key properties of curiosity that any theory should seek to address: (1) its intensity, (2) transience and dependence on immediate stimulus, (3) association with impulsivity, and (4) tendency to disappoint when satisfied. In our formulation, curiosity arises from such a gap, defined by the comparison between the amount of simplification occurring at a moment in time and the amount of simplification that is deemed to be immediately possible. Curiosity may, of course, be unsatisfied if we are unable to perform the relevant action (e.g., clicking the link, asking a person a key question, peering to see what someone is reading, and so on), in which case we would expect the curiosity to be replaced with, or at least mixed with, frustration.

Loss aversion, in our model, helps to explain curiosity's intensity; loss aversion means that people will be especially strongly motivated to engage in sense-making when they perceive a gap between how much sense-making they have achieved and how much they believe could be achieved.

Curiosity's transience and dependence on immediate stimuli, in our model, can be explained by the transience of expectations of sense-making (which serves as the reference level in our model). The classic example of curiosity's transience is when one is walking behind someone and becomes curious of what they look like from the front; as soon as they turn the corner and disappear, curiosity tends to vanish almost instantly. Such a pattern is well explained by an account in which the individual's disappearance virtually eliminates the potential to obtain the missing perspective, thus altering expectations and curiosity.

Curiosity's association with impulsivity is easily addressed by the proposal that sense-making is a drive, similar to hunger, thirst and sexual desire, all of which are commonly associated with impulsive behavior. Drives tend to naturally increase one's preference for drive-related consumption (indeed, that is their very function), but future drive levels are uncertain. So immediate drives tend to promote immediate consumption, and the drive for sense-making is no exception, except that the 'consumption' in its case, involves information acquisition.

Finally, the tendency to disappoint when satisfied is also explained by loss aversion, which amplifies an individual's motivation to obtain information ("wanting," to use terminology proposed by Berridge, 1996) without necessarily having a commensurate impact on pleasure from obtaining the information ("liking," in Berridge's terminology).

4.2. Boredom

Boredom, in our framework, arises from the absence of sense-making and the comparison of this paucity of sense-making to some higher baseline point of reference the negative feeling associated with the inability to engage in sense-making. Boredom has two elements. First, we have the feeling that there are no actions, Q, that we can take that are likely to lead to information, a, that will help us make sense of any information, K, that we are interested in. So, when considering actions, Q, which might lead to such data, we suspect that $\left\langle S_R(e(K|B \cup \{a\}))\right\rangle_Q$ is low. This might arise if, for example, we are in a waiting room stocked with magazines detailing the lives of celebrities we care nothing about; if we are stuck a hotel with only the prospect of futile channel-hopping; if we come to a halt in a railway tunnel and all we can scan from the window is the reflection of our empty carriage; or if we stuck in an unchallenging and repetitive job, from which we are learning nothing new.

There is a second element of boredom: that there is no prospect of shifting our attention to some new K which we can make sense of. Thus, we can sometime stave off boredom, by viewing our environment in a different way: e.g., wondering whether celebrity magazines may be telling us something about the human condition; comparing the news agenda from different channels; puzzling about the explanation of the double-reflections in carriage windows; or thinking about last night's exciting TV drama during a day of repetitive work. But attempting to shift K is effortful, and may be impossible if, for example, on a conference call we need to pay attention to what is said. In short, then, we are bored when our drive for sense making is thwarted, and there is nothing we can do, either by our actions, a, or by changing the 'topic' K, to find anything that we both want to, and are able to, make sense of.

This account of boredom is closely related to, but distinct from, a recent "opportunity cost" account of boredom which sees boredom as closely related to the feeling of mental effort (Kurzban et al., 2013). In this alternative account, the "sensation

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

of 'mental effort' is the output of mechanisms designed to measure the opportunity cost of engaging in the current mental task. According to this alternative account the key to whether an activity is boring (or mentally effortful) is whether the task in question is seen as valuable.

Our analysis of boredom is different from the opportunity cost account. First, our account does not view boredom as related to a negative feeling of mental effort. According to our account, we can be bored while experiencing no feeling of mental effort—indeed, this is typically the case when our context and task is under-stimulating or repetitive. Conversely, we can feel a strong sense of mental effort, e.g., when solving a cross-word puzzle or Sudoku, even though the task clearly has no extrinsic 'value' and is a purely an entertainment. Indeed, were this not the case, such effortful pastimes would presumably universally be perceived as boring, and their vast popularity would be mysterious. Moreover, our account also does not predict that the feeling that a task is valuable will decrease one's boredom with it—so a task such a detecting very rare signals of enemy aircraft on a radar monitor may be recognized as vitally important but also as dreadfully boring. Instead, the key determinant of boredom according to our account is the (in)ability to engage in sense-making.

Both accounts can make sense of some salient 'facts' about boredom. One is that tasks that require minimal amounts of attention can be more boring than tasks that don't require any attention (e.g., staring at a blank wall). Both theories can account for this observation by noting that tasks which don't require attention allow for mind-wandering (in our account, actively changing K), to something that can be very much associated with sense-making—e.g., thinking about a story, a puzzle or one's own life. Another interesting fact about boredom is that people rarely report being bored when trying to go to sleep, even unsuccessfully—this is especially puzzling, given that attempts at sleep typically involve minimal sensory stimulation. Our account of boredom would explain this observation by noting that expectations about sense-making, and hence the appropriate reference points, are likely to be very low in this situation; the period before sleep is not typically a time of substantial sense-making, probably by evolutionary design, so there is no expectation that sense will be made and hence no unfavorable comparison with actual sense-making.

4.3. Flow, and the short-circuiting of sense-making

'Flow' (Csikszentmihalyi, 1990) refers to a state of concentration or complete absorption with an activity one is engaged in. Flow is the antithesis of both boredom and curiosity, in its determinants, consequences, and in the feelings it evokes. Contrary to boredom, flow is associated with high levels of sense-making, and contrary to curiosity, there is no craving for missing information. In our account, expected sense making is high; and such expectations are typically achieved.

Flow is generally seen as a good thing, but people can, and often do, experience flow when engaged in activities that are difficult to construe as beneficial to them. Most drives can be effectively short-circuited by products and activities that activate, and appear to satisfy, the drive while providing little sustained benefit. The concept of 'empty calories' from fast food that, through the introduction of salt and fat, appeals to our evolutionary programming but yields little real nutritional benefit is a paradigmatic example; internet pornography is another. Modern electronic game makers provide the analog to these examples when it comes to sense-making. Games like "Angry Birds" give the brain the perception of continual sense-making, even though the sense that is made fails to add up to anything useful. As we indicated above, the popularity of Sudoku and crosswords, as well as other pastimes including jigsaw puzzles, word- and picture-search, spot-the-difference problems, and logic problems, as well as domestic 'organizing' activities without obvious practical function, such as gardening, stamp and coin collecting, and many art and craft leisure activities, suggests that creating an organized environment (and hence an environment which we can readily make sense of) is intrinsically rewarding.

4.4. Confirmation bias

Piaget (1954) drew a contrast between two ways of responding to information: assimilation, which involves making sense of the information in light of existing explanations of the world, and accommodation, which involves adapting such explanations to take account of new, incongruent, information. Relatedly, Kuhn, in the *Structure of Scientific Revolution* (1970), proposed that science doesn't progress in a continuous fashion, but in discontinuous jumps: in essence, when assimilation of new data to the existing framework is replaced by accommodating new data by the creation of a new theoretical framework. But, both in cognitive development and in science, shifting to a new framework—i.e., accommodating rather than merely assimilating—is not undertaken lightly. Indeed, there can even be active resistance to the need for a new framework—as captured in the physicist Max Planck's famous quip that "Science advances one funeral at a time."

'Confirmation bias' (e.g., Nickerson, 1998) is closely related to these phenomena. It refers to the tendency for people to seek out and interpret information in a fashion that tends to support existing beliefs, rather than requiring those beliefs to be updated. Confirmation bias is well documented experimentally and in real-world settings, but its origins and motivational underpinnings have been much less studied.

One reason that the brain is likely to stick with existing explanations of knowledge, K, is simply that finding alternative better models is difficult. According to our approach, we can only entertain one explanation, e, of information K, at a time; and the natural way to search for better explanations, e', is by making local adjustments to the current explanation. Indeed, most models of learning in the brain and cognitive sciences involve a highly local search for better 'models' (e.g., using gradient descent or some stochastic generalization, Dayan and Abbott, 2001; Rumelhart and McClelland, 1986). So if the brain locally searches the space of explanations, then we should anticipate a degree of inertia or 'stickiness' with regard to existing

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

explanations. For this reason, we should expect people to become, to some extent, "stuck" in particular interpretations of an ambiguous image (e.g., the Rubin face-vase stimulus of Fig. 1), yet occasionally flipping their interpretation abruptly.

While important, the 'inertia through local search' style of explanation is far from a complete account. In particular, it does not explain the active resistance to new explanations (and, indeed, active dismissal of apparently awkward data) mentioned above. Our account provides an explanation of this motivation. When considering a particular set of data K, a new explanation e' may appear superior to our current explanation e. But if e' has wide application to other aspects of our life, then by accepting e' we may be committed to rethinking may other things, K', K'', K'''. And before we do so, we do not know whether these aspects of our life will make us feel better about ourselves or worse about ourselves; or whether our lives will make more sense, or less sense. To the extent that losses loom larger than gains in the context both of valuations, V, and sense-making S, (captured in our model by the fact that changes in both with respect to the reference point are modulated by prospect-theory style functions μ_V and μ_S), we should expect people to be averse to such re-thinking: the fear of the result of rethinking having negative consequences is likely to outweigh the possibility that it may have positive consequences.

We suspect that the sense-making term, S, may play a key role for two reasons. First, people are often resistant to new knowledge, and consequent explanatory change in domains which appear irrelevant to their evaluation of their own lives—e.g., while acting as jurors in a court case (Nickerson, 1998). Second, people typically experience a period of disorientation before existing facts are fully assimilated into the new explanation, during which they have a feeling that the facts, or even their lives, no longer make sense. Such temporary loss of 'sense-making' is, as we would expect, often highly aversive. Atherton (1999, p. 88), in a paper on "resistance to learning," drew a distinction between "additive" and 'supplantive' learning (similar to Piaget's between assimilation and accommodation) and suggested that resistance to the latter is greater because of its accompanying element of loss:

"The depression and confusion experienced by people experiencing supplantive learning follows a very similar pattern to that of those passing through crises [. . .]. Having been de-stabilised, they pass through a period of *disorientation*, from which they emerge to *re-orientation*."

4.5. Information-avoidance

Although closely related to confirmation bias, information avoidance is worthy of separate treatment because our model provides a novel account of how and why it occurs. Information avoidance presents a kind of paradox because, at some level, to avoid information selectively one has to have at least some idea about the content of the information. If expectations are unbiased, moreover, then positive surprises should be as likely as negative surprises. So one might suppose that the prospect of good or bad news would balance out and that there would be no particular reason to avoid information.

Different accounts of information avoidance have been proposed. One account, based on disappointment aversion (Kőszegi, 2003, 2010) which is closely related to loss aversion and reference-dependence in our model), posits that negative surprises are more unpleasant than positive surprises are pleasant—our account embodies this insight as we shall see below. Another (Oster et al., 2013) assumes that people adopt optimistic expectations, but are aware at some level that they are doing so, and so avoid information to protect themselves against having their unrealistic expectations get shattered. Yet a third (Karlsson et al., 2009) proposes that knowing something definitively has a greater impact on utility than simply suspecting something; when information is expected to be adverse, this 'impact effect' motivates information avoidance.[6]

Our theoretical perspective provides a related but somewhat different interpretation from these existing accounts. Note, first that the autonomy of the information process system, in finding the simplest explanation of the data which it is currently processing—i.e., maximizing the sense-making term, S, implies that we cannot *directly* influence our construal of our lives to improve the value term V. We are not able, for example, simply able to willingly misperceive, i.e., to misread, our teaching ratings as a collection of 9s and 10s, when they are actually 2s and 3s. Similarly, we are not able willfully to tells ourselves, falsely, that 1 represents the 'top' of the scale and '10' the bottom; or alternatively that most of our colleagues will be rated as straight 1s, so that 2s and 3s represent a strong performance. The autonomy of the sense-making system requires that we know bad news when we see it, which is likely of crucial evolutionary importance: Were we able to shape our perception of the world to maximize V, not S, then rather than creating a model of the world that made the most of the available information, we would create a 'happy dream' unconnected to the external world.

But, though the deliberative system cannot influence the autonomous system directly, it can do so indirectly—by affecting what information the autonomous system has available to process. So while the 'digestive processes' of the sense-making system may be autonomous, and outside the influence of the aim of optimizing value, its 'diet' is not. In particular, we can decide to avoid sources of information that we suspect may lead us to bad evaluations V of how our lives are going; and, given the severe attentional limits of the cognitive system, we can also selectively decide what information to pay attention to from the vast amount of information available to us. So, while we cannot decide to misread our teaching ratings in order to maintain our positive mood and preserve a positive self-image, we can decide not to look at the ratings further; or, if we have glanced at them, and noticed an alarming number of apparently low scores, we can avert our eyes, put them into the recycling bin, and vow inwardly to think no more about them.

[6] See Golman et al. (2015) for a review of evidence, strategies, and theoretical accounts of information avoidance.

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

For similar reasons, we may also avoid sources of information that may threaten to disturb our current explanations. Typically, of course, learning more information, if it is relevant at all, is likely to help us understand the world better (e.g., think about the hint about the alphabet in Fig. 2). Although it is possible that new information may undermine our current understanding (e.g., learning that the Fig. 2 comes from Ancient Egypt, so that the Latin alphabet cannot help explain the pattern), the receipt of information that decreases sense-making is likely fairly rare. Indeed, there is a well-known theorem in the mathematical theory of coding, the data-processing inequality (Cover and Thomas, 1991), which states that the *expected* code-length of a set of data, in the light of information, a, from any observation or experiment, A, cannot be greater than the initial code-length:

$$C(K|B) \geq \left\langle C(K|B \cup \{a\}) \right\rangle_{a \in A} \tag{11}$$

If we assume that sense-making is merely the inverse of code-length (i.e., f in Eq. (3) is the identity function), then this would imply that any experiment or observation will, in expectation, improve sense-making. Note, though, that, as with valuation, any change in sense-making is assessed using a prospect-theory style value function, μ_S, in which losses loom larger than gains.

For this reason, it is entirely possible that, for some possible observations or experiments, A:

$$\left\langle \mu_S(S)K|B)) > \mu_S(S(K|B \cup \{a\})) \right\rangle_{a \in A} \tag{12}$$

That is, the expected utility from a change in sense-making may be negative for some experiments or observations—and hence these are sources of information that we will be motivated to avoid. An academic who has developed a particular theory, for example, might stop running critical experiments that could challenge the theory. Presumably, if the academic truly believes the theory, new data should be much more likely to support the theory than to refute it; however the down-side of the latter could easily more than outweigh the more probable but smaller upside of the former.

In this type of case, avoiding subjecting one's theory to a really rigorous test may be motivated both by the desire to avoid both a reduction in sense-making and a reduction in one's evaluation of one's life—if, for example, one's sense of self-worth is connected to the validity of the theory. In more common situations, however, these two motives are likely to be pitted against one-another. Generally, important new information will tend to simplify the encoding of current information, as we have seen Eq. (11). So, other things being equal, one might expect that gathering new information will generally improve sense-making. But the anticipated impact of new information on value, in contrast, has a neutral expected value

$$\left\langle V(e(K|B \cup \{a\})) - V(e(K|B)) \right\rangle_{a \in A} = 0 \tag{13}$$

From the point of view of absolute value, one should feel entirely indifferent about whether to obtain new information—losses and gains will balance out. But, because losses loom larger than gains, the expectation of V_R term will typically be negative (although the curvature of the value function, μ_V will also matter—e.g., there may be a positive expected evaluation where for example, there is a high probability of mildly good news, and a small probability of very bad news. So, for example, one might want to know the results of an exam where one is almost sure one has passed).

The choice of whether to look or not at an information source, therefore, is likely to depend on the balance between the typically positive expectation of the S_R term and the typically negative expectation of the V_R term. Whether we choose to look at our teaching ratings, investigate whether our spouse is having an affair, or whether our child is engaged in illegal activities, will depend on whether curiosity overcomes trepidation, and the most common conflicted pattern will be to seek information that we strongly suspect will make us miserable (Kruger and Evans, 2009).

There is, of course, a further impetus to sample information: Learning new information may help to inform decision making and allow us to change our actions in beneficial ways. Although new information can, by ill luck, sometimes lead us to switch to worse decision options, the expected quality of decisions should never fall after the receipt of new information (for a relevant formal result, see Juslin et al., 2006). In general, therefore, psychological factors that lead us to avoid information will tend to have a negative impact on decision quality. For example, suspecting that her teaching ratings are bad, a teacher may not look at them; and therefore not get valuable feedback about how to improve (i.e., not be able to change her actions appropriately); and hence become locked into a spiral of poor teaching. Similarly, not taking a medical test may mean that a person is not treated early enough to halt the progress of a serious disease, and so on.

In these situations, we should expect information avoidance to be especially prevalent among people who discount the future in an extreme, or hyperbolic, fashion, putting disproportionate weight on the present. Present-bias would lead to a disproportionate weighting of the short-term costs of looking as compared with the long-term benefits of gaining the feared information.

4.6. Consumer choice

In a recent experiment, Ellen Evers and colleagues (Evers et al., 2014a; see, earlier, Evers et al., 2014b) presented college students with a choice between a set of cheap plastic pens, each a different color, and a notebook. Other students were presented with the same choice, but, added to the cheap plastic pens were two more pens of the same color as one of the pens in the prior set. A majority of students chose the smaller set of all-different-colored pens over the notebook; however, despite the fact that the pen collection was objectively enhanced by the addition of the two same-colored pens, a majority of

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

students chose the notebook over the larger collection. This pattern violates monotonicity, which is one of the simplest and widely accepted principles of good decision making. It can be understood as a manifestation of the preference for simplicity: "all different" is much simpler to encode than the mish-mash of same and different colors that was created by the addition of the two same-color pens.

Evers and colleagues showed, too, showed that how items were organized made a big difference to choice behavior. For example, when the two pens in the above example were separated from the other pens, creating one all-different set and another all-same set, then preferences shifted back in favor of the pens over the notebook. Similarly, when people chose sets of bottles of beer, they preferred those with a simple organization, either being all different or all the same; a set of beers which mostly differ but have some repetitions was less preferred.

As a body, these experiments show the people have a strong preference for sets of items that can be organized simply—in our terms, they are easy to make sense of. Indeed, the authors further attribute these preferences to the preference for sets that are easily described—i.e., that have a short code-length. A final study in the paper sought to provide strong evidence that the preference for all-different and all-same was, in fact, derivative of the preference for descriptive simplicity. In this study students made choices between two box-sets of BB King compact disks, which were deliberately chosen to be about comparable in desirability. One of the two box sets, however, was randomly selected for each subject to be described as all having been recorded in Biloxi, or Tupelo, two locations which a pretest had identified as *not* particularly desirable locations for a recording to have been made. Despite the undesirability of these recording locations, providing a description that, in effect, 'made sense' of one collection or the other tended to increase preference for that collection.

4.7. Esthetics

The question of why some pieces of art, movies, plays, literature etc. are widely appreciated whereas others are not has received surprisingly little attention from social and behavioral scientists, perhaps because it is seen as a problem too difficult to yield workable insights. Indeed, as represented most vividly by the famous paper by Stigler and Becker (1977) "De gustibus non est disputandum" (there is no arguing with tastes), many economists have been adamant that the origin of tastes is a topic that lies outside of the purview of their profession. Although certainly not addressing all relevant causal factors (e.g., there is certainly a huge amount of social 'herding' incumbent in art appreciation), our framework does provide some hints about determinants of esthetic tastes.

Quite analogous to our discussion of two ways that the drive for sense-making plays into curiosity (i.e., the desire to make sense of stimuli provided, such as riddles, and the desire for information that promises to make sense of existing, stored, information), one can imagine two ways that sense-making plays into esthetics.

The first is that we are likely to have a preference for materials that can be made sense of—books and movies with coherent plots; representational art and so on. The desire for sense-making within the stimuli provided is probably most characteristic of what is sometimes referred to as "lowbrow" art—art that is enjoyable to experience but doesn't leave us with much in the long-term.

The second is the preference for literature, art and so forth that may not provide much sense-making in and of itself, but helps us to make sense of other aspects of our lives or our world. The painting Guernica, for example, is somewhat painful to look at, but gives the viewer a new perspective on war. As Schmidhuber (2009, p. 11) notes in a paper that is closely related to this one, "good observer-dependent art deepens the observer's insights about this world or possible worlds, unveiling previously unknown regularities., connecting previously disconnected patterns in an initially surprising way."

Even though the second category can help to explain why much art is anything but simple to make sense of, it probably doesn't go sufficiently far in explaining the preference for such work. Indeed, there almost seems to be some pleasure, or at least excitement, to be derived from art that *doesn't* make sense of the world, or that even defies sense-making. In a letter to his brothers, the poet John Keats used the term "negative capability" to refer, in laudatory terms, to a situation in which "a man is capable of being in uncertainties, mysteries, doubts, without any irritable reaching after fact and reason." (Bate, 2009, p. 249) He coined the term (and its definition) as a contrast to the proclivities of the poet Coleridge, who, Keats complained to his brothers, was "incapable of remaining content with half-knowledge." Coleridge, according to Keats, was searching for a single, higher-order truth or solution to the mysteries of the natural world, a task which Keats, who saw the world as infinite and impenetrable, viewed as crass.

The pleasures of esthetics need not be limited to the traditional arts. Sunstein (2015), for example, in an insightful book review titled "How Star Wars Illuminates Constitutional Law," writes about the pleasures of learning, in the climactic scene of The Empire Strikes Back, that Darth Vader is Luke Skywalker's father. As Chris Taylor (2014) expresses it in the book that Sunstein reviews, this suddenly explains "at a strike why everyone from Uncle Owen to Obi-Wan to Yoda has been so concerned about Luke's development, and whether he would grow up to be like his father." Sunstein argues that similar moments happen in constitutional legal scholarship: "In constitutional law, many law professors argue for such moments, for example by recognizing new limits on the power of the federal government or new rights of various sorts."

4.8. Science

Science is sometimes viewed as the antithesis of art, yet our framework points to some key commonalities. Scientists are certainly motivated, in part, by a desire the search for fame and fortune, but many if not most are probably also driven by a

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

motive more closely related to the current perspective—the desire for sense-making. As Glynn (2010) points out in his book "Elegance in Science," scientists as well as lay-people can be thrilled by their assimilation of a new theoretical perspective that sheds light on disparate, otherwise confusing, facts. Wilkinson (2015), likewise, in an article on mathematician Yitang Zhang's solution of the "bound gaps" problem in math, writes that:

> Pure mathematics, as opposed to applied mathematics, is done with no practical purposes in mind. It is as close to art and philosophy as it is to engineering. "My result is useless for industry," Zhang said. The British mathematician G. H. Hardy wrote in 1940 that mathematics is, of "all the arts and sciences, the most austere and the most remote." Bertrand Russell called it a refuge from "the dreary exile of the actual world." Hardy believed emphatically in the precise aesthetics of math. A mathematical proof, such as Zhang produced, "should resemble a simple and clear-cut constellation," he wrote, "not a scattered cluster in the Milky Way." Edward Frenkel, a math professor at the University of California, Berkeley, says Zhang's proof has "a renaissance beauty," meaning that though it is deeply complex, its outlines are easily apprehended. The pursuit of beauty in pure mathematics is a tenet. Last year, neuroscientists in Great Britain discovered that the same part of the brain that is activated by art and music was activated in the brains of mathematicians when they looked at math they regarded as beautiful.

We reproduce the passage in full because it touches on so many of the themes discussed in this paper. First, the fact that the "result is useless for industry" highlights the fact that insight is valued in its own right, apart from any material gains it confers. Second, Hardy's view that a mathematical proof should resemble a simple and clear-cut constellation" as opposed to a "scattered cluster" makes clear that mathematical proofs are very much a matter of sense-making, as does Wilkinson's description of Zhang's proof as "deeply complex" but with "outlines [that] are easily apprehended." Frenkel's comment that Zhang's proof has "a renaissance beauty" draws connections between scientific insight and esthetic appreciation, as does the reference to neuroscience research.

The pleasure derived from sense-making is not only a source of scientific progress, but also of stagnation, because once a scientist has arrived at a particular form of sense-making, and especially if they are attached to it in virtue of having proposed it or adopted it over a span of time, they are likely to become resistant to abandoning it—for reasons closely related to those discussed in the subsection above on confirmation bias.

4.9. Concern about others' beliefs

One of the most consequential, yet rarely studied, features of human motivation is the extent to which people care about others' beliefs. People sort geographically on the basis of beliefs, avoid interacting with people holding different beliefs and, when they can't avoid interacting with people holding different beliefs, avoid the 'conversational mine-fields' associated with actually discussing their differences. People expose themselves to media consistent with their existing beliefs, proselytize to try to bring others' beliefs into conformity with their own and, at the extreme, seek to silence those they disagree with, in some cases by trying to kill them.

Why do people care about others' beliefs; why don't we take a live-and-let-live attitude toward what is, after all, invisible in other people's minds? Our theory of sense-making provides one possible explanation. Beliefs, according to the theoretical perspective propounded here, are part and parcel of sense-making; people choose beliefs that fit with the larger sense that they make of the world. A negative change in the degree to which one's life makes sense is aversive, and this is especially true as a result of the inclusion of the change term, and is further exacerbated by loss aversion. Confronting beliefs that are incompatible with one's own, perhaps via an interpersonal encounter or exposure to media, according to this perspective, is aversive because it is impossible to escape the recognition that (at least) one set of beliefs must be wrong, and it may not be possible to confidently rule out the possibility that it is one's own beliefs, and the sense-making they are part of, that are deficient. (See Golman et al., 2015, for a related, and more fully developed perspective on this issue.)

4.10. Conspiracy theories and religion

A "conspiracy theory" (a term coined by Karl Popper) is, according to Wikipedia, "an explanatory proposition that accuses two or more persons, a group, or an organization of having caused or covered up, through secret planning and deliberate action, an illegal or harmful event or situation." Although conspiracies undeniably do occur, the term has taken on a pejorative connotation (Moore, 2002) due to the widespread perception that conspiracies are suspected to occur far more often than they actually do, as well as to the unlikely nature (e.g., implausible degree of coordination between disparate characters) of many conjectured conspiracies. Social scientists have found that conspiracy theories are widespread, and that conspiracy theories are especially embraced by the poor, minorities, and those who exhibit low levels of trust and high levels of economic insecurity (Goertzel, 1994).

A wide range of explanations for the pervasiveness of conspiracy theories have been proposed, but a very large number of these involve discomfort with uncertainty and lack of understanding of events. Popper himself argued that conspiracy theories deny the reality that most important events have multiple causes and are the product of a multitude of agents; conspiracy theories reflect the implicit assumption that significant events must have been intended and caused by some individual or group. Sunstein and Vermeule (2009), in one of the most comprehensive treatments of

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

the subject, argue that the appeal of many conspiracy theories "lies in the attribution of otherwise inexplicable events to intentional action, and to an unwillingness to accept the possibility that significant adverse consequences may be a product of invisible hand mechanisms (such as market forces or evolutionary pressures) or of simple chance, rather than of anyone's plans" (p. 6). The authors then continue that "people do not like to believe that significant events were caused by bad (or good) luck, and much prefer simpler causal stories. In particular, human "minds protest against chaos," and people seek to extract a meaning from a bewildering event or situation, a meaning that a conspiracy may well supply."

Interestingly, the explanations that have been proposed for the attraction of conspiracy theories may also help to explain the pervasiveness of belief in God. Park (2005) argues, and supports with correlation research, the idea that traumatic events strengthen belief in God because of the threat they pose to non-randomness. In a study of 169 students who had experienced the death of a significant other in the past year, Park found that, in the short-run, those who were religious suffered a greater drop in well-being following the death, presumably because their misfortune challenged their belief in a benign power. However, in the long-term—i.e., for those who had experienced the death further in the past—the effect reversed, and religiosity was associated with better coping and higher subjective well-being. Kay et al. (2010) also argue that attempts to cope with perceptions of randomness may be a key factor in religious beliefs. "Affirming the existence of a controlling God," the authors write, "may provide an excellent means for insulating oneself from the aversive arousal associated with randomness." (p. 216) Providing some, albeit weak, experimental evidence for this proposition, the authors found that subjects who completed a word unscrambling exercise which exposed them to words related to randomness (e.g., "chance," "random") as opposed to negative words (e.g., "slimy") but who were not subject to a misattribution manipulation, reported greater belief in supernatural control.

As should already be evident, both conspiracy theories and religion can potentially be viewed as manifestations of the drive for sense-making. However, they represent quite different approaches. Although Sunstein and Vermeule posit a preference for "simpler causal stories," many if not most conspiracy theories are actually fantastically complex. For example, the popular conspiracy theory that the U.S. government was behind the 9–11 tragedy envisions the government hiring the hijackers, booby-trapping the buildings (according to the dominant theory, the airplanes alone would not have been sufficient to cause their collapse), and so on. In the language of information theory, this is a very complicated theory ($C(e'|B)$ is high) that makes perfect sense of the world ($C(K|e', B)$ is low). Religious beliefs are, in some sense, the opposite; they are a very simple theory: "god causes everything" ($C(e'|B)$ is low) that makes sense of, in the sense of predicts, almost nothing, leaving $C(K|e', B)$ largely unchanged.

4.11. The importance of narrative

Another manifestation of the drive for sense-making is, likely, the human affinity for narrative—for telling stories about things that help to make sense of them. Pennington and Hastie (e.g., 1991, 1992) find that jurors are much more persuaded by narratives, which make sense of the facts, than by logical arguments. Pennebaker finds in numerous studies (see, e.g., Pennebaker and Seagal, 1999) that writing about an emotional personal experience for as little as 15 min a day confers mental and physical benefits in as little as three days. "The act of constructing stories. . ." Pennebaker and Seagal write, "allows one to organize and remember events in a coherent fashion. [and] gives individuals a sense of predictability and control over their lives." (p. 1243) Alice Gregory (2015), in an article about a text-message based counseling service, writes that "people who spent their high-school years chatting with friends on landlines are often dismissive of texting, as if it might be a phase one outgrows, but the form is unparalleled in its ability to relay information concisely. The act of writing, even if the product consists of only a hundred and forty characters composed with one's thumbs, forces a kind of real-time distillation of emotional chaos."

Jerome Bruner (in his early career a pioneer of research on sense-making) focused, in late career, on narrative. Bruner, like Pennebaker, views the construction of stories as a natural human process that helps people to make sense of their lives and ultimately shapes how they live. As Bruner writes, "The self-telling of life narratives achieves the power to structure perceptual experience, to organize memory, to segment and purpose-build the very 'events' of a life. In the end, we become the autobiographical narratives by which we 'tell about' our lives." (2004, p. 694) Bruner's point is that people don't only try to make sense of their life by forming a narrative of it, but organize their life not only to conform to the narrative they create, but with an eye toward burnishing it and maintaining cohesion.

Less momentously, the drive for sense-making might play a role in the human propensity to 'organize' (Glushko, 2013). A typical office or living room has, for example, books relatively neatly arranged on shelves, rather than in piles on the floor. And, although many academic offices violate these norms, visitors' negative reactions could be interpreted as providing further support for the power of the desire for organization. Of course, some aspects of organizing clearly serve pragmatic functions (e.g., so that things can be found easily, or use as little storage space as possible). But many aspects of organizing (e.g., aligning pictures) have no obvious function; and the existence of extreme organizing behaviors—compulsions—which can significantly impede, rather than contributing to the achievement of a person's wider goals, strongly suggest that the desire to organize also has a deeper basis. A curious observation, for which we have no explanation (but suspect that the explanation is interesting) is that procrastination—avoiding doing unpleasant tasks that one should be devoting oneself to—so often results in organizing as a substitute activity.

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

4.12. The image and 'the good life'

Early in the second half of the twentieth century, two very different books were published with the same title, The Image, published in 1956 by Kenneth Boulding, an economist, discussed how images, and the desire to conform to them, guide human social behavior. Boulding discusses many of the basic ideas central to this paper. For example, he discusses how popular concepts shape our vision of reality. He notes that "Veblen, for instance, was not, I think, a great social scientist, and yet he invented an undying phrase: "Conspicuous consumption." After reading Veblen, one can never quite see a university campus or an elaborate house in just the same light as before" (Boulding, 1956, p. 9). Boulding also talks about how beliefs can change dramatically as a result of a kind of 'regime shift,' which he refers to a "change of the image which might be described as a revolutionary change." Boulding attributes the suddenness of the change to "the fact that our image is in itself resistant to change. When it receives messages which conflict with it, its first impulse is to reject them as in some sense untrue." Daniel Boorstin, a historian, published his The Image 6 years later (in 1962), a book which presented more of a social critique of how symbol (image) had replaced substance in American culture. Boorstin describes currents in American culture whereby the reporting of an event becomes more "real" than the event itself. Boorstin coined the term "pseudo-event" to describe events or activities that serve little purpose other than to be reported via advertisements or media. Boorstin would undoubtedly have felt vindicated if he were alive today to witness the culture of twitter and 'selfies'.

Whatever one's value judgments, imagery has almost certainly played a key role in human life from early in the emergence of the species. Most people walk around with an image of their current situation if not an image of the world they inhabit. And, if we are to judge from the ubiquity of song lyrics touting different conceptions of the good life (songlyrics.com produces 224 songs with good life *in the title*) most people also probably have an image of how they would like their lives to be. Such images of the good life probably play an important role in decision making, simplifying otherwise impossibly difficult decisions. When it comes to momentous decisions such as whether and who to marry, where to live, what to study, it is extremely difficult to calculate costs and benefits. It is much easier, however, to make a judgment about whether a particular partner, career or location takes one closer, or propels one farther, from one's image of the good life.

5. Conclusion

In a characteristically insightful paper titled "The Mind as a Consuming Organ," Thomas Schelling (1987) drew attention to how much human, and economic, activity is devoted to consumption, not of conventional 'goods and services,' but of events happening purely in the mind. As Schelling wrote, "we consume with our mouths and noses and ears and eyes and proprioceptors and skin and fingertips and with the nerves that react to external stimuli and internal hormones; we consume relief from pain and fatigue, itching and thirst. But we also consume by thinking. We consume past events that we can bring up from memory, future events that we can believe will happen, contemporary circumstances not physically present like the respect of our colleagues and the affection of our neighbors and the health of our children" (p. 354). The consumption-from-thinking that Schelling wrote about was all about pleasures and pains of memory and imagination, and about self-esteem—the value side of our model. Our central point in this paper is that such belief-based utility still leaves out an extraordinarily important form of mind-consumption: People also care tremendously about, and devote substantial time, money and attentional resources, to gaining insight, acquiring knowledge, and making sense of the world.

Most globally, people care not only about how good their lives are, but also about *meaning*. An earlier paper, on "The Economics of Meaning" (Karlsson et al., 2004) proposed four possible interpretations of what meaning entails: (1) meaning as a resolution of preferences; (2), meaning as an extension of oneself either socially or temporally; (3) meaning as an act of making sense of one's life; and (4) meaning as an assertion of free will. Our focus in this paper is on the third of these conceptions of meaning. We argue that much information collection and processing occurs, not in an effort to make or perceive our lives in a more positive fashion, but to make sense of our world and of our lives.

In Section III of this paper we proposed a simple model of the drive for sense-making. The model draws a distinction between autonomous information processing that occurs without conscious volition and external processes that are at least partly driven by volitional actions. The model assumes that the two processes are characterized by different objective functions. Internal processing aims only at maximal simplification of a given body of information. We assume further that the output of the process is a single construal (interpretation) of the information at any point in time. External processes differ from internal processes on a number of dimensions. First, the external processes don't engage directly in sense-making; their function instead is to select the information provided to the internal process. Second, external processes reflect two underlying motivations: (1) the desire for sense-making (simplification) but also (2) the desire to construe the world in favorable ways.

Subsequent sections of the paper then showed how this simple model can make sense of a wide range of phenomena, specifically: curiosity, boredom, flow, confirmation bias, information avoidance, common patterns of consumer preference, esthetic preferences, patterns of progress and stagnation in science, belief in God and conspiracy theories, and the role of narrative and imagery in daily life.

Although we believe that the model proposed in the paper sheds considerable light on each of these applications, all models are simplifications, and we suspect that the current one leaves out some potentially important phenomena. One such phenomenon was discussed in the section on esthetics, in connection with Keats' notion of negative capability: Clearly, there are situations in which people value stimuli that don't make sense—that are impossible to simplify—precisely for

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

that characteristic. Another, not discussed in the paper, is the pleasure one can sometimes get from a change in the sense one makes of the world, even when the new sense is no simpler than the old sense. As long as one isn't too invested in one interpretation of one's knowledge, a sudden transformation to a radically different interpretation can actually be exceptionally pleasurable, producing, as Sunstein (2015, p. 5) describes it, "an 'aha', a shiver, a tingle of the spine."

The motives discussed in this paper are by no means new but have been central to human existence from our beginnings. Indeed, they may be present in other species such as cats, which are anecdotally notable for their curiosity. However, informational motivations are especially prominent in humans, and appear to be assuming an ever more important role in human affairs. As numerous commentators have pointed out (e.g., Gleick, 2011), we are well into what could be called the 'information age'—a time in which information has become much more central to our lives and economies than the more tangible products that played a more central role in earlier periods.

Simplification has, of course, an upside and downside. On the one hand, simplification brings tremendous efficiencies in decision making and problem solving. On the other hand, simplification is associated with destructive stereotyping, dogged attachment to outmoded theories, the pursuit of images of the good life that probably wouldn't bring us much pleasure were we ever to attain them and to the substitution of symbol for substance in American, and probably world, culture. Whatever its ups and downs, identifying the existence of the drive for simplification, and its role in daily judgment and decision making, will hopefully help social scientists to make sense of the complexities of human existence.

Acknowledgments

We thank Peter Dayan and Cass Sunstein for comments on earlier versions of these ideas, and seminar participants at Harvard and Concordia Universities for helpful comments. Chater was supported by ERC grant 295917-RATIONALITY, the ESRC Network for Integrated Behavioural Science [grant number ES/K002201/1], the Leverhulme Trust [grant number RP2012-V-022], and Research Councils UK Grant EP/K039830/1. Loewenstein thanks the STICERD at the London School of Economics for hosting him during the sabbatical year when the collaboration was begun. Thanks also to Nikolos Gurney and Rosa Stipanovic for help preparing the manuscript.

References

Atherton, J., 1999. Resistance to learning: a discussion based on participants in in-service professional training programmes. J. Vocation. Educ. Train. 51 (1), 77–90.
Bate, W.J., 2009. John Keats. Harvard University Press, Cambridge, MA.
Baumeister, R.F., Bratslavsky, E., Finkenauer, C., Vohs, K.D., 2001. Bad is stronger than good. Rev. Gen. Psychol. 5 (4), 323–370.
Berridge, K.C., 1996. Food reward: brain substrates of wanting and liking. Neurosci. Biobehav. Rev. 20 (1), 1–25.
Blakemore, C., 1993. Vision: Coding and Efficiency. Cambridge University Press, Cambridge.
Boorstin, D.J., 1962. The Image, or, What Happened to the American Dream. Vintage.
Boulding, K., 1956. The Image: Knowledge in Life and Society. University of Michigan Press.
Bruner, J., 2004. Life as narrative. Soc. Res. 71 (3), 691–710.
Chaitin, G.J., 1995. The Berry paradox. Complexity 1 (1), 26–30.
Chater, N., 1996. Reconciling simplicity and likelihood principles in perceptual organization. Psychol. Rev. 103 (3), 566.
Chater, N., 1999. The search for simplicity: a fundamental cognitive principle? Q. J. Exp. Psychol. 52 (A), 273–302.
Chater, N., 2005. A minimum description length principle for perception. In: Pitt, M., Myung, I. (Eds.), Advances in Minimum Description Length: Theory and Applications. MIT Press, Cambridge, MA.
Chater, N., Vitányi, P., 2003. The generalized universal law of generalization. J. Math. Psychol. 47, 346–369.
Chater, N., Vitányi, P., 2007. 'Ideal learning' of natural language: positive results about learning from positive evidence. J. Math. Psychol. 51, 135–163.
Clark, H.H., 1996. Using Language. Cambridge University Press, Cambridge.
Cover, T.M., Thomas, J.A., 1991. Elements of Information Theory. John Wiley & Sons.
Csikszentmihalyi, M., 1990. Flow: The Psychology of Optimal Experience. Harper and Row, New York.
Dayan, P., Abbott, L.F., 2001. Theoretical Neuroscience. MIT Press, Cambridge, MA.
Evers, E.R.K., Inbar, Y., Loewenstein, G., Zeelenberg, M., 2014a. Order Preference, Available at SSRN: http://ssrn.com/abstract=2466991.
Evers, E.R.K., Inbar, Y., Zeelenberg, M., 2014b. Set-fit effects in choice. J. Exp. Psychol.: Gen. 143 (2), 504–509.
Fodor, J.A., 1983. Modularity of Mind. MIT Press, Cambridge, MA.
Fukushima, K., 2001. Recognition of partly occluded patterns: a neural network model. Biol. Cybern. 84 (4), 251–259.
Gleick, J., 2011. The Information: A History, a Theory and a Flood. HarperCollins.
Glushko, R.J., 2013. The Discipline of Organizing. MIT Press, Cambridge, MA.
Glynn, I., 2010. Elegance in Science: The Beauty of Simplicity. Oxford University Press, Oxford.
Goertzel, T., 1994. Belief in conspiracy theories. Polit. Psychol. 15, 731–742.
Golman, R., Haggman, D., Loewenstein, G., Working Paper 2015. Information Avoidance. Department of Social and Decision Sciences, Carnegie Mellon University.
Gregory, A., 2015, February 9. R U There?: A New Counselling Service Harnesses the Power of the Text Message.
Hardy, G.H., 1940. A Mathematician's Apology. Cambridge University Press, Cambridge.
Hochberg, J., McAlister, E., 1953. A quantitative approach to figure "goodness". J. Exp. Psychol. 46, 361–364.
Hoffman, D.D., Richards, W.A., 1984. Parts of recognition. Cognition 18 (1), 65–96.
Hsee, C.K., Ruan, B., Working Paper 2014. Increasing Happiness by First Inducing and Then Resolving a Desire: The Case of Curiosity. The University of Chicago.
Idesawa, M., 1991. Perception of 3-D illusory surface with binocular viewing. Jpn. J. Appl. Phys. 30 (4B), L751.
Juslin, P., Fiedler, K., Chater, N., 2006. Less is more in covariation detection—or is it. In: Fiedler, K., Juslin, P. (Eds.), Information Sampling and Adaptive Cognition. Cambridge University Press, Cambridge, UK, pp. 92–123.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–291.
Kanizsa, G., 1979. Organization in Vision: Essays on Gestalt Perception. Praeger Publishers.
Karlsson, N., Loewenstein, G., McCafferty, J., 2004. The economics of meaning. Nordic J. Polit. Econ. 30 (1), 61–75.
Karlsson, N., Loewenstein, F., Seppi, D., 2009. The ostrich effect: selective avoidance of information. J. Risk Uncertainty 38 (2), 95–115.

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016

Kay, A.C., Moscovitch, D.A., Laurin, K., 2010. Randomness, attributions of arousal, and belief in God. Psychol. Sci. 21 (2), 216–218.
Kemeny, J.G., 1953. The use of simplicity in induction. Philos. Rev. 62, 391–408.
Koffka, K., 2013/1935. Principles of Gestalt Psychology. Routledge.
Kőszegi, B., 2003. Health anxiety and patient behavior. J. Health Econ. 22 (6), 1073–1084.
Kőszegi, B., 2010. Utility from anticipation and personal equilibrium. Econ. Theory 44 (3), 415–444.
Kőszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. Q. J. Econ. 121 (4), 1133–1165.
Kruger, J., Evans, M., 2009. The paradox of Alypius and the pursuit of unwanted information. J. Exp. Soc. Psychol. 45 (6), 1173–1179.
Kuhn, T.S., 1962. The Structure of Scientific Revolutions, 1st ed. University of Chicago Press, Chicago.
Kurzban, R., Duckworth, A., Kable, J.W., Myers, J., 2013. An opportunity cost model of subjective effort and task performance. Behav. Brain Sci. 36 (06), 661–679.
Levinson, S., 2000. Presumptive Meanings. MIT Press, Cambridge, MA.
Li, M., Vitányi, P.M., 2009. An Introduction to Kolmogorov Complexity and Its Applications. Springer Science & Business Media.
Loewenstein, G., 1994. The psychology of curiosity: a review and reinterpretation. Psychol. Bull. 116 (1), 75–98.
Moore, R., 2002. Reconstructing reality: conspiracy theories about Jonestown. J. Popular Cult. 36 (2), 200–220.
Nickerson, R.S., 1998. Confirmation bias: a ubiquitous phenomenon in many guises. Rev. Gen. Psychol. 2 (2), 175–220.
Oster, E., Shoulson, I., Dorsey, E.R., 2013. Optimal expectations and limited medical testing: evidence from Huntington disease. Am. Econ. Rev. 103 (2), 804–830.
Park, C.L., 2005. Religion as a meaning-making framework in coping with life stress. J. Soc. Issues 61 (4), 707–729.
Pennebaker, J.W., Seagal, J.D., 1999. Forming a story: the health benefits of narrative. J. Clin. Psychol. 55 (10), 1243–1254.
Pennington, N., Hastie, R., 1991. Cognitive theory of juror decision making: the story model, A. Cardozo Law Rev. 13, 519–557.
Pennington, N., Hastie, R., 1992. Explaining the evidence: tests of the story model for juror decision making. J. Pers. Soc. Psychol. 62 (2), 189–206.
Piaget, J., 1954. The Construction of Reality in the Child. Basic Books, New York.
Pomerantz, J.R., Kubovy, M., 1986. Theoretical approaches to perceptual organization: simplicity and likelihood principles. In: Boff, K.R., Kaufman, L., Thomas, J.E. (Eds.), Handbook of Perception and Human Performance: Vol. H. Cognitive Processes and Performance, 36. Wiley, New York, pp. 1–45.
Pylyshyn, Z., 1999. Is vision continuous with cognition?: the case for cognitive impenetrability of visual perception. Behav. Brain Sci. 22 (03), 341–365.
Rissanen, J., 1987. Stochastic complexity. J. R. Stat. Soc. Ser. B 49, 223–239.
Rock, I., Palmer, S., 1990. Gestalt psychology. Sci. Am. 263, 84–90.
Rumelhart, D.E., McClelland, J.L., 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, 2 vols. MIT Press, Cambridge, MA.
Scheffler, I., 1972. Vision and revolution: a postscript on Kuhn. Philos. Sci. 39, 366–374.
Schelling, T.C., 1987. The mind as a consuming organ. In: Elster, J. (Ed.), The Multiple Self. Cambridge University Press, Cambridge UK, pp. 177–195.
Schmidhuber, J., 2009. Simple algorithmic theory of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes. J. Soc. Instrum. Control Eng. 48 (1), 21–32.
Stevenson, A. (Ed.), 2010. Oxford Dictionary of English. Oxford University Press, Oxford.
Stigler, G.J., Becker, G.S., 1977. De gustibus non est disputandum. Am. Econ. Rev. 67, 76–90.
Sunstein, C.R., 2015. How star wars illuminates constitutional law. Mich. Law Rev. (forthcoming).
Sunstein, C.R., Vermeule, A., 2009. Conspiracy theories: causes and cures. J. Polit. Philos. 17 (2), 202–227.
Taylor, C., 2014. How Star Wars Conquered the Universe. Head of Zeus, London.
Wallace, C.S., Freeman, P.R., 1987. Estimation and inference by compact coding. J. R. Stat. Soc. Ser. B 49, 240–251.
Wilkinson, A., 2015. The pursuit of beauty: Yitang Zhang solves a pure-math mystery. New Yorker February 2.
Wright, E., 1992. Gestalt switching: Hanson, Aronson, and Harré. Philos. Sci. 59, 480–486.

Please cite this article in press as: Chater, N., Loewenstein, G., The under-appreciated drive for sense-making. J. Econ. Behav. Organ. (2015), http://dx.doi.org/10.1016/j.jebo.2015.10.016